

## Mathematical Model of Opinion Dynamics in Social Groups

**Evgeny Gorbatiykov**

*Keldysh Institute of Applied Mathematics Russian Academy of Science*  
E-mail: gor-104@mail.ru

**Elena Kornilina**

*Keldysh Institute of Applied Mathematics Russian Academy of Science*  
E-mail: ekornilina@gmail.com

**Alexander Mikhailov**

*Keldysh Institute of Applied Mathematics Russian Academy of Science*  
E-mail: apmikhailov@yandex.ru

**Alexandr Petrov**

*Keldysh Institute of Applied Mathematics Russian Academy of Science*  
E-mail: petrov.alexander.p@yandex.ru

Doi:10.5901/mjss.2013.v4n10p380

### Abstract

Mathematical models describing relationships in a group of individuals are known for a long time. The basic idea underlying these models can be described as follows. There are affective (emotionally colored) not necessarily symmetrical relationships between any two members of the group. In general, these relationships vary over time and can be described by numerical functions of time. The value of the corresponding function is positive if the first individual has a positive attitude toward the second one and negative if the attitude is negative. While interacting, two individuals share their views on other members of the group. It affects the dynamics of group relationships. At the same time, the model uses principles such as "if the person whom I dislike says something unpleasant about a third person then I improve my opinion of this third person." This paper also deals with the modeling of group interactions, with the difference that the considered interactions have features of mass (not paired) communications, i.e., the opinion of each individual is equally available to all the others (at the same time, affective relations are paired). This situation occurs, for example, in communities of users of internet forums and blogging platforms. Another important difference is due to the fact that individuals do not discuss other members of the group, but discuss various issues on the topic (e.g., political) of interest to all members of the group. The results of numerical experiments for the system are presented and a number of substantial conclusions are formulated.

**Keywords:** mathematical modeling; group dynamics; political positions; pre-election programs.

### 1. Introduction

This paper deals with the development of a dynamic model of proximity positions of participants in closed groups who express their opinion on given set of questions.

Consider a group with a fixed number of participants (i.e., closed), discussing a number of issues in a way that everyone's opinion is immediately available to everyone else. The position of the participants may change over time under the influence of statements of other members of the group. There're affective relationships between participants (one may be positive or negative to the other), which, in particular, influence individuals to change their positions on the issues. These relationships may also change over time.

An example of such interactions could be discussions on forums, online communities, etc. Due to the significant development of the popularity of the exchange of positions in the Internet, it becomes particularly important to study the interactions of this type, using the methods of mathematical modeling.

Mathematical models of group dynamics are studied for at least 70 years of the twentieth century, they deal with the processes of changing individual positions under the influence of the discussions within a closed group. We present several simulation results of steady (static) group relations. According to the theory of structural balance proposed by F. Heider, a member of the group will be in a state of cognitive balance if his relations in each of the triad to which he belongs to have one of the following forms: a friend of a friend is my friend (R1), the foe of my foe is my friend (R2), a friend of my foe is my foe (R3), the foe of my friend is my foe (R4). From the Heider's theory, the stable triads are those in which all participants are close to each other, or two close to each other and are far from the third. In Heider-Cartwright-Harary Theorem, it's shown that in the group in which the participants share their opinions (positive or negative) pairwise, the steady state leads to the partition into two stable subgroups.

The J. Davis's balance is a generalization of the concept of the previous balance. It is based on an assumption that the triad will also be stable in the case when all the participants are foes. Such approach allows to model situations with the several antagonistic sub-groups that are in a stable condition.

For the model proposed in this paper, a priori, it was not obvious what type of balance it is implemented, more specific or more general. The study revealed that the model belongs to the more general type (the above-cited model of Davis), which allows to implement a wider class of stable states.

The closest dynamic models to the proposed one are the models of J. Hunter, V. Kovchegov, P. Killworth-G. Bernard. We can also mention the work of D.A. Gubanov, D.A. Novikov, A.G. Chkhartishvili, which deals with the dissemination of information in social networks.

The main feature of these study is that we consider not only the pairwise relations. The opinion of each individual's available to everyone else – these are the mass communications' features. At the same time, each of the individual has his own opinion (positive or negative) to another, so that there is the type for group modeling – mass-pairwise communication.

The second part of the paper is devoted to obtaining data for the saturation of model. The main values that underpin the model – are the relationship between the individuals (subjective distance at which one perceives the other), and their positions on each of the discussed issues.

The measurement of the relationship between individuals in groups could be found in empirical work on sociometry, for example, articles by J. Moreno and later works. In the present study, the new approaches to the measurement of the relationship between individuals are not offered.

There are a number of methods to measure the proximity of positions in case if participants discuss political issues only. For example, there are lots of methods based on the analysis of voting results (see, e.g. various American methods.). Speaking on the analysis of the positions expressed in political texts, we can state that the traditional methods are based on the use of variations on the content analysis, which requires expert evaluation. In particular, one of the most well-known methods of analysis of political programs is represented in the project "Manifesto", its basic element is a special classifier. There are other methods based on content analysis.

It is natural to start developing the way of non-expert measuring the position/proximity. The method suggested in this paper based on the analysis of political texts using latent semantic analysis (LSA). The applying this method became possible because of the proposed hypothesis that the texts that express similar positions (e.g., political) will be syntagmatically close. Syntagmatic proximity, in turn, is the basis of the LSA.

Purpose - to the construction and study of mathematical models of the dynamics near a closed group of participants' positions and openly express their opinions on a fixed set of questions.

## 2. Construction and study of the mathematical model

We consider a closed group of individuals  $K$ , discussing  $N$  issues. We assume that  $x_{ij}(t) \geq 0$  characterizes the subjective distance of the individual  $i$ th to the  $j$ th;  $z_i^n(t) \in \mathbb{R}$  – position of the individual  $i$ th on  $n$ th issue at time  $t$ . The variable  $x_{ij}(t)$  should be understood an aggregated value that characterizes the subjective distance at which one participant perceives himself to the other, while, in general,  $x_{ij}(t) \neq x_{ji}(t)$ . We assume that the subjective distance of a participant in relation to himself is zero:  $x_{ii}(t) = 0$ ,  $i = 1, \dots, K$ .

The position  $z_i^n(t)$  can be understood as a positive or negative attitude of a participant towards to the discussed problem.

According to the condition of the model, the participants express their positions publicly, which leads to the changing of relationship between them and their positions afterward. Our model will be based on the following assumptions about the connection between the subjective distances and positions:

1. the position on the  $n$ th issue at time  $t$  are considered to be close enough to improve relations between the participants, when  $|z_i^n(t) - z_j^n(t)| < a$ , where  $a$  is some given positive parameter;
2. if it is true for most of the discussed issues then the subjective distance decreases:  $dx_{ij}(t) / dt < 0$ ;
3. the positions on the issues of the  $i$ th participant are attracted to positions of  $k$ th participant if they are within a zone of its perception as a like-minded person:  $x_{ik} < R_i$ , where  $R_i > 0$ ;
4. the influence of other participants are added together.

These assumptions are written as the following equations of the model: (1)-(3)

$$\frac{dx_{ij}}{dt} = \gamma_i x_{ij}(t) \sum_{n=1}^N \alpha_i^n \left[ (z_i^n(t) - z_j^n(t))^2 - a^2 \right] \quad (1)$$

$$\frac{dz_i^n}{dt} = \sigma_i \sum_{j=1}^K \beta_{ij}^n (z_j^n(t) - z_i^n(t)) (R_i - x_{ij}(t)) \quad (2)$$

$$x_{ij}(0) = x_{0ij}, \quad z_i^n(0) = z_{0i}^n \quad (3)$$

$i, j = 1, \dots, K, i \neq j, n = 1, \dots, N.$

Parameter  $\alpha_i^n$  describes the subjective importance of the  $n$ th issue for the  $i$ th individual,  $\sum_{n=1}^N \alpha_i^n = 1$ , and the parameter  $\gamma_i \geq 0$  characterizes the individual in terms of how quickly or slowly he is changing his attitude to other members of the group.

The parameter  $\sigma_i \geq 0$  describes the propensity of an individual to change his positions, and the parameter  $R_i > 0$  describes how close participants should be to perceive them as like-minded people. The impact on the position of the  $i$ th individual on the  $n$ th issue of all the other individuals are summarized with coefficients  $\beta_{ij}^n \geq 0$ , normalized so that  $\sum_{j=1}^K \beta_{ij}^n = 1$ .

Thus, the system of equations modeling the dynamics of positions has the form (1), (2) and represents a Cauchy problem with initial conditions (3). Generally, the system has  $K(K-1) + KN$  equations.

We introduce an additional character. Let  $y_i^n(t) = z_i^n(t) - z_1^n(t)$ ,  $i = 2, \dots, K, n = 1, \dots, N$ . Then the system (1)-(2) can be rewritten as follows:

$$\frac{dx_{1j}}{dt} = \gamma_1 x_{1j} \sum_{n=1}^N \alpha_1^n \left[ (y_j^n)^2 - a^2 \right], j = 2, \dots, K \quad (4)$$

$$\frac{dx_{i1}}{dt} = \gamma_i x_{i1} \sum_{n=1}^N \alpha_i^n \left[ (y_i^n)^2 - a^2 \right], i = 2, \dots, K \quad (5)$$

$$\frac{dx_{ij}}{dt} = \gamma_i x_{ij} \sum_{n=1}^N \alpha_i^n \left[ (y_i^n - y_j^n)^2 - a^2 \right], i, j = 2, \dots, K \quad (6)$$

$$\begin{aligned} \frac{dy_i^n}{dt} = & -\sigma_i \beta_{i1}^n y_i^n (R_i - x_{i1}) \\ & + \sigma_i \sum_{j=2}^K \beta_{ij}^n (y_j^n - y_i^n) (R_i - x_{ij}) - \sigma_1 \sum_{j=2}^K \beta_{1j}^n (y_j^n) (R_1 - x_{1j}), \end{aligned} \quad (7)$$

$i = 2, \dots, K, n = 1, \dots, N$

Definition 1. Define consent state of (4)-(7) its steady state solution where  $x_{ij} = x_{ji} = 0$ ,  $i, j = 1, \dots, K$ ,  $y_i^n = 0$ ,  $i = 2, \dots, K$ ,  $n = 1, \dots, N$ .

The definition should be understood as follows: in the consent state everyone believes the other is a like-minded, and they have the same position on all issues.

Theorem 1 (the stability of the steady state). Suppose that the system (4)-(7) is so that all the coefficients  $\alpha_i^n, \beta_{ij}^n, \gamma_i, \sigma_i, a, R_i$ , where  $i, j = 1, \dots, K$ ,  $n = 1, \dots, N$ , are positive. Then the consent state of the system is asymptotically stable.

Consider the problem

$$\left\{ \begin{array}{l} \frac{dx_1}{dt} = \gamma_1 x_1 [y^2 - a^2] \\ \frac{dx_2}{dt} = \gamma_2 x_2 [y^2 - a^2] \\ \frac{dy}{dt} = \sigma y [x_1 + x_2 - 2R] \\ x_1(0) = x_{10}, x_2(0) = x_{20}, y(0) = y_0 \end{array} \right. \quad \begin{array}{l} (8) \\ (9) \\ (10) \\ (11) \end{array}$$

where  $y = z_1^1 - z_2^1$  and  $x_{10} > R, x_{20} > R, y_0 > a$ .

Theorem 2 (of disagreement). Let the coefficients  $\gamma_1, \gamma_2, \sigma, a, R$  be positive and let  $T > 0$  – be a number such that the solution of the Cauchy problem (8)-(11) exists in the interval  $0 \leq t \leq T$ . Then the inequalities

$$\begin{array}{l} x_1(t) > \alpha_1(t), \alpha_1(t) = x_1(0) \exp[\gamma_1(y_2(0) - a^2)t] \\ x_2(t) > \alpha_2(t), \alpha_2(t) = x_2(0) \exp[\gamma_2(y_2(0) - a^2)t] \\ y(t) > \alpha_3(t), \alpha_3(t) = y(0) \exp[\sigma(x_1(0) + x_2(0) - 2R)t] \end{array}$$

are true for all  $t: 0 < t \leq T$ .

The meaning of Theorem 2 can be explained as follows: in the conditions of the theorem, positions and relations of participants to each other diverge. The divergence is interpreted as a failure for participants to have similar positions in the given system if the number of variables is fixed (the participants can get close only if a new element will be added to the system, e.g. a new participant).

Numerical simulation were produced using the Runge-Kutta RK4 methods of (1)-(3) for different initial conditions.

1. It is shown that if the number of participants ( $K$ ) = 3, and the number of discussed issues ( $N$ ) = 1, then the following stable states are realized: all the distances between the participants are close to each other, or between two are close, but not to the third one (see Fig. 1).

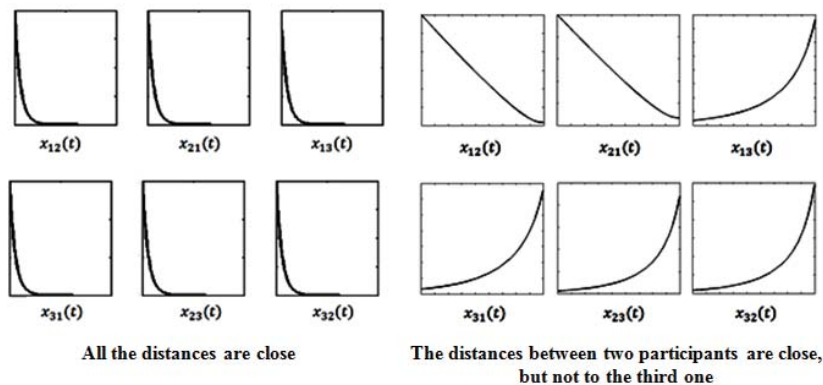


Fig. 1. Stable states  $K = 3, N = 1$

2. If  $K = 4, N = 1$ , the following stable states are realized: all distances between the participants are close, the distances between the three are close, the distances of the two pairs are close, the distance is close only in one pair.

Thus, it's shown that the model corresponds to the traditional models of group dynamics (eventually, the relations between participants are established in a way of not making dissonance) and the possibility of realizing more than two stable states is demonstrated (as in the approach of J. Davis).

### 3. Measuring the proximity of political positions

Note that the system (1)-(3), which describes the discussed model, has the values of two types: 1)  $x_{ij}(t)$  – the subjective distance from  $i$ th participant to the  $j$ th 2) the difference  $z_i^n(t) - z_j^n(t)$  – the position's proximity between the  $i$ th and  $j$ th participants on the  $n$ th issue.

The first value  $x_{ij}(t)$  is measured by sociometric and psychological methods in a fairly well-developed science area that is being developed since the XIX century.

The second value  $z_i^n(t) - z_j^n(t)$  is traditionally measured by methods based on content analysis, which requires the use of expert judgment (for example, in the project "Manifesto"). The paper proposes an alternative method of determining the proximity of political positions expressed in the texts.

The technique consists in determining the proximity of positions by analyzing syntagmatic proximity of texts that express these positions. Syntagmatic approach to measuring distances in the text is understanding the meaning of words through the context in which it is used. In this case, the distance is measured by comparing syntagmatic properties. This approach has been implemented in the method of latent semantic analysis (LSA), the patented method proposed by American scientists. LSA is based on "the hypothesis that between the individual words and generalized context (sentences, paragraphs and whole texts) in which they are found, there is an implicit (latent) relationship, causing a set of mutual constraints." The underlying assumption is that these relationships are different in the texts, expressing different political positions. So, it is easy to imagine that the name of the politician is generally used in a positive way by his supporters, and in the negative by his opponents. This applies not only to the names and the names of the parties, but also to certain political events, projects, etc. Thereby expressing the political position of the text can be classified by highlighting the context in which the individual words are immersed in these texts.

The basis for the application of the method of latent semantic analysis is the proposed hypothesis:

Hypothesis. The proximity of political positions is related to the syntagmatic proximity of the texts (fragments forming the texts), expressing these positions.

In contrast to the traditional method of using LSA, the additional preprocessing of the texts is conducted – lemmatization and removal of specially selected stop words (words that do not have the semantic weight).

#### 4. Experiments on measurement the proximity of political positions

A software package that uses the proposed empirical method of non-expert numerical measurements of political positions' proximity has been developed. By means of the developed software package pre-election programs of political parties in the elections of the 2007th and the 2011th years in the State Duma of the Federal Assembly of Russia have been analyzed.

Comparisons were made for the four parties: Edinaja Rossija, the Communist Party, the Liberal Democratic Party, Yabloko. Three of them are traditionally represented in the State Duma, we have also included the program of Yabloko, as the party has a reputation as the "most unlike" to the other three. This expected "otherness" was the reason that a special part of the paper is the comparison of Edinaja Rossija and Yabloko. Pre-election programs (except these four) of other parties may also be included in the review, the method does not contain any restrictions on the number of party programs studied.

One way to present the results of the analysis is to construct a diagram similar to those shown in Fig. 2. Each chunk corresponds to one row and one column of the diagram. A cell located at the intersection, for example, the 10th row and 25th column (the same as cell 25th row and 10th column) describes the syntagmatic proximity between 10th and 25th chunks. At the same time, the more syntagmatically close are the two chunks, the darker will be the cell.

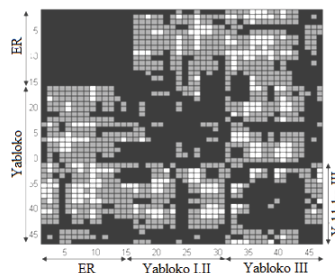


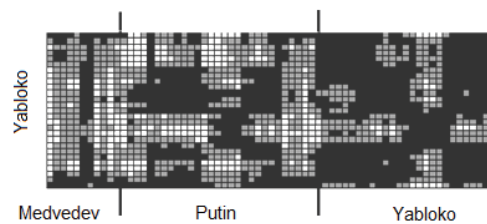
Fig. 2. Diagram of similarity for the programs Edinaja Rossija and Yabloko in 2007

On Fig. 2 chunks 1-15 belong to the program of Edinaja Rossija (ER), chunks 16-46 – the program of the party "Yabloko". Dark square, made up of rows and columns 1-15, shows a high syntagmatic proximity of the chunks and in

accordance with the proposed hypothesis – the proximity of political positions expressed in these chunks. The diagram also shows that the program of Yabloko distinguishes two parts, each of which has a higher internal connection than the program as a whole. The boundary between them is on the 30th chunk. A more detailed analysis has shown that it corresponds to the boundary of sections II "The alternative – the welfare state" and III "The well-being for everyone" of Yabloko program.

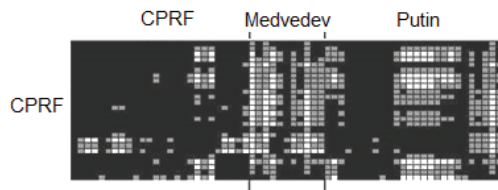
The ER program (chunks 1-15) is clearly separated from the Yabloko program in which there are two internally connected parts, relevant to sections I, II (16-30 chunks), and section III (31-46 chunks).

At the congress of Edinaja Rossija it was decided to consider the texts of Dmitry Medvedev's and Vladimir Putin's speeches to be the electoral program in the 2011th. A priori, one would assume that Medvedev's speech, who has the reputation of a liberal, would be closer to the Yabloko program than Putin's speech. However, latent semantic analysis shows (Fig. 3 shows only the rows that match Yabloko program) this statement not to be true: the intersection of columns 1-11, related to Medvedev's text and rows 41-67, related to the Yabloko party has a lighter tone than the intersection of columns related to the text of Putin and the lines related to the text of Yabloko party.



**Fig. 3.** Similarity diagram of Edinaja Rossija and Yabloko pre-election programs in 2011 (only the rows that match the Yabloko program are shown)

The comparison of Communist Party of Russian Federation program (CPRF) and ER program in the 2011th shows in particular that they are poorly distinguishable if we take ER program as a whole. If we consider (Fig. 4) the correlation of lines 1-26 (CPRF) separately to columns 27-37 (Medvedev) and 38-64 (Putin), we can see the dissimilarity of the program of the Communist Party with Medvedev's speech, high affinity with the first part and the dissimilarity with the last part of Putin's speech.



**Fig. 4.** Similarity diagram of CPRF and Edinaja Rossija pre-election programs in 2011 (only the rows that match the CPRF program are shown)

We denote the syntagmatic distance between the  $i$ th and  $j$ th chunks as  $r_{ij}$  and define

$$A = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_1} \left( \frac{1+r_{ij}}{2} \right) + \sum_{i=N_1+1}^{N_1+N_2} \sum_{j=N_1+1}^{N_1+N_2} \left( \frac{1+r_{ij}}{2} \right)}{N_1^2 + N_2^2} \quad (12),$$

$$B = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=N_1+1}^{N_1+N_2} \left( \frac{1+r_{ij}}{2} \right) \quad (13).$$

Definition 2. A measure of the syntagmatic proximity of programs is the value  $R = 200 B / A$ .

Thus calculated values R for the pre-election programs of four Russian parties in the 2007 and 2011 elections are shown in the Table 1 and Table 2.

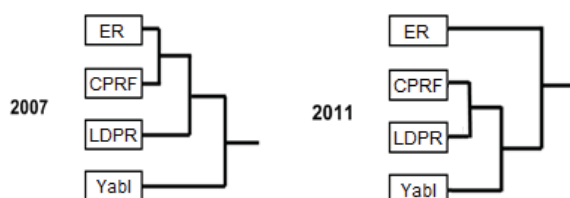
**Tab. 1.** Syntagmatic proximity of pre-election programs in 2007

| 2007    | ER | CPRF | LDPR | Yabloko |
|---------|----|------|------|---------|
| ER      | 0  | 99   | 92   | 87      |
| CPRF    | 99 | 0    | 95   | 91      |
| LDPR    | 92 | 95   | 0    | 91      |
| Yabloko | 87 | 91   | 91   | 0       |

**Tab. 2.** Syntagmatic proximity of pre-election programs in 2007

| 2011    | ER | CPRF | LDPR | Yabloko |
|---------|----|------|------|---------|
| ER      | 0  | 87   | 91   | 89      |
| CPRF    | 87 | 0    | 96   | 85      |
| LDPR    | 91 | 96   | 0    | 90      |
| Yabloko | 89 | 85   | 90   | 0       |

Using the method of hierarchical clustering an analysis of data in Table 1, 2 has been made: which programs are the closest and the most distant. The results are shown in Fig. 5.



**Fig. 5.** The structure of syntagmatic proximity of programs in 2007 и 2011

Thus, in the pre-election campaign in 2007 the closest programs were the programs of ER and the Communist Party, and the most distant from the others was the Yabloko program. In the campaign of 2011 was a certain restructuring: the closest programs were the programs of Communist Party and the Liberal Democratic Party, and the most distant from the others was ER program. Overall, the 2011 programs of the different parties have become more similar to one another than in 2007.

## 5. Acknowledgement

Made with the support of RFBR project №№ 12-01-31461-mol\_a, 13-01-00392-a

## References

- Cartwright D., Harary F. Structural Balance: A Generalization of Heider's Theory. - Psych. Rev., 63, 1956, p. 277-293
- Davis J.A. Clustering and Structural Balance in Graphs // Hum. Relat. 1967 V.20 p.181-188.
- Davis J.A., Leinhardt, S. The Structure of Positive Interpersonal Relations in Small Groups // Sociological theories in progress. V.2 Boston, 1972, p.218-251.
- Heider F. Attitudes and Cognitive Organization. - J. of Phych., 21, 1946, p. 107-112
- Hunter J.E. Dynamic sociometry // J. Math. Sociol. 1978. V.6. p.87-138.
- Kemeny J.G., Snell J.L. Mathematical Models in the Social Sciences. - New York: Blaisdell Publishing Co., 1962 reprinted by M.I.T. Press, Cambridge, Mass, 1972
- Killworth P.D., Bernard H.R. A Model Of Human Group Dynamics // Social Science Research. Vol. 5. 1976. P. 173-224
- Landauer T., P.W. Foltz, D. Laham. Introduction to Latent Semantic Analysis. Discourse Processes 25: 259-284 (1998).

- Laver M., Benoit K., Garry J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2), 311–331.
- Lowe W., Benoit K., Mikhaylov S., Laver M. Scaling Policy Preferences From Coded Political Texts. *Legislative Studies Quarterly* 26(1, Feb): 123-155.
- Martin L.W., Vanberg G., A robust transformation procedure for interpreting political text, *Political Analysis* 16 (1), 93-100.
- Moreno, J. L.. *Sociometry, Experimental Method and the Science of Society. An Approach to a New Political Orientation*. Beacon House, Beacon, New York – 1951.
- Segalovich I. A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine. *MLMTA'2003*, Las Vegas
- Volkens A., O.Lacewell, P.Lehmann, S.Regel, H.Schultze, A.Werner (2011): The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR), Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).
- Благовещенский Н.Ю. Оценка политических позиций в Государственной Думе 3-го созыва по результатам голосований. М.: Фонд ИНДЕМ, 2004.
- Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. «Социальные сети: модели информационного влияния, управления и противоборства», 2010 – 228 стр.
- Ковчegov В.Б. Модель динамики групповых структур человеческих сообществ // *Социология* 4М. 1991. № 1. С. 75-98.
- Митрофанова О.А. Семантические расстояния: проблемы и перспективы // XXXIV Международная филологическая конференция: Вып. 21. Прикладная и математическая лингвистика. СПб., 2005.